U.S. Grid Testbed Status and Plans

Kaushik De

Univ. of Texas at Arlington

ATLAS Software Week

CERN, May 2002

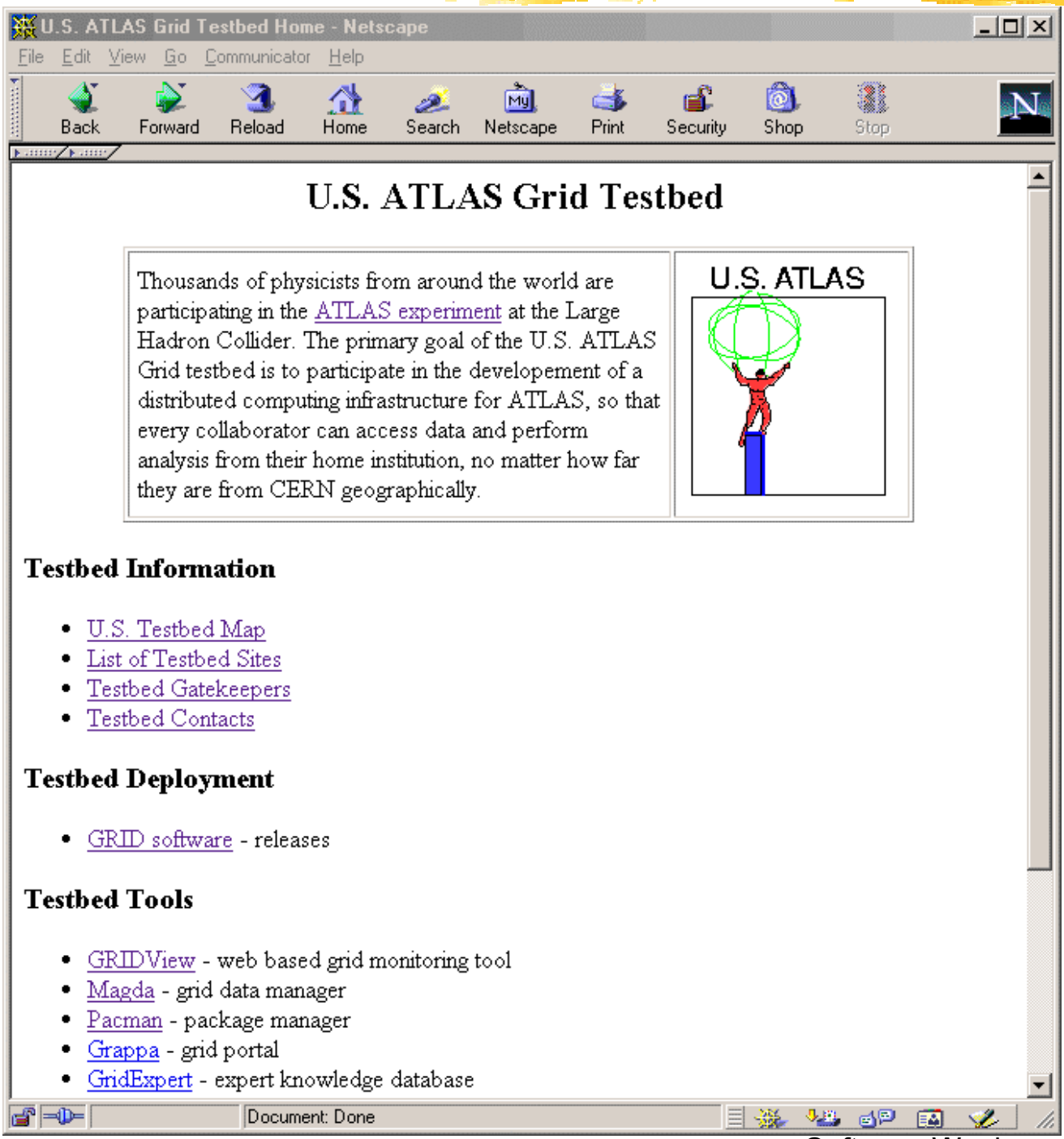
Outline

- Many recent US Grid Testbed meetings:
  UTA - April, BNL - May, BU - June, LBNL - July
  http://heppc1.uta.edu/atlas/workshop_april_2002/index.html
  http://www.usatlas.bnl.gov/computing/software/core-grid-200205/
- Report on evolving plans
  - Testbed status
  - Software distribution
  - Application toolkit
  - MC production plans
  - Integration
  - Grid tools
  - SC2002 demos

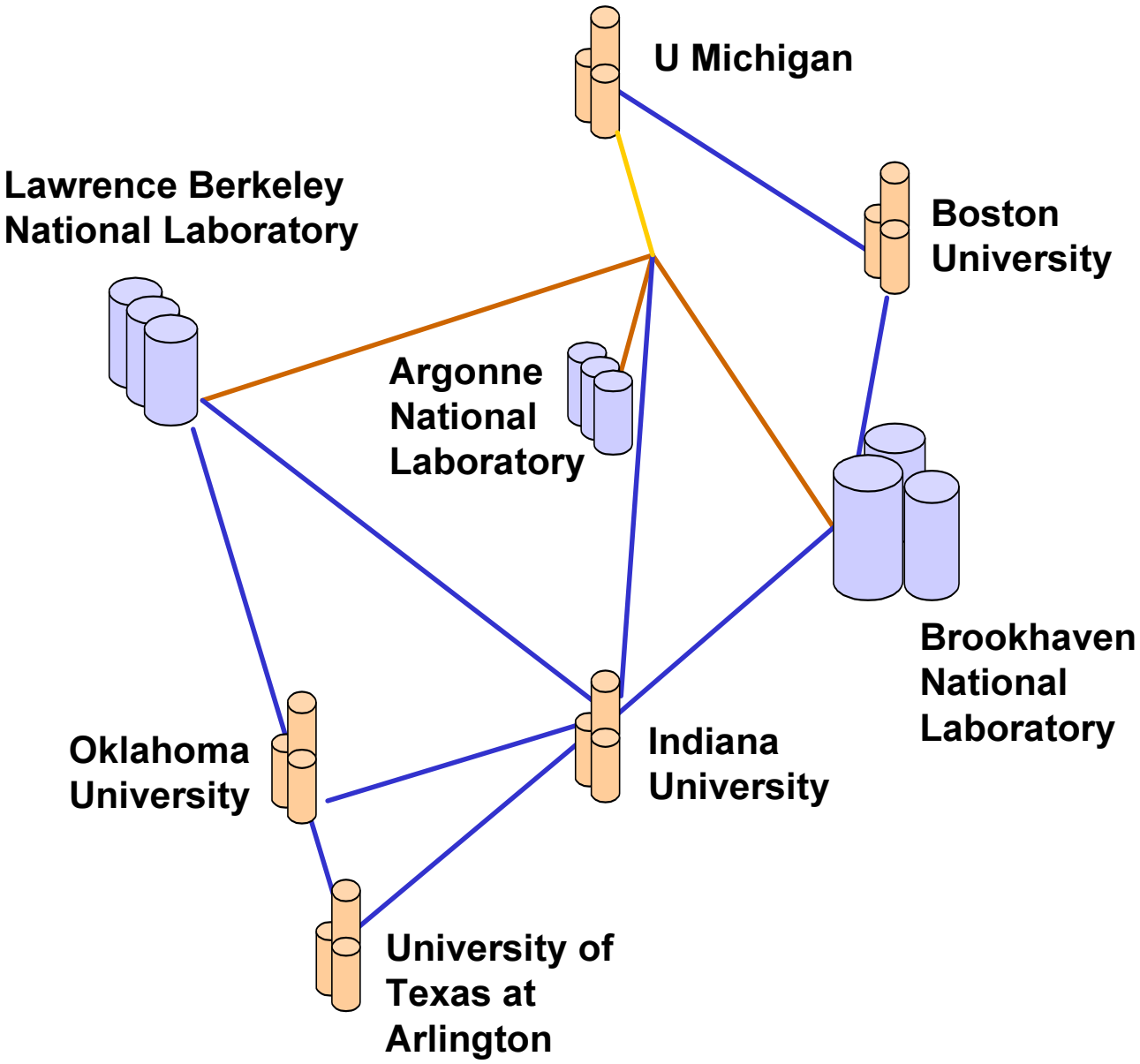
Testbed Goals

- Demonstrate success of grid computing model for HEP
  - in data production
  - in data access
  - in data analysis
- Develop & deploy grid middleware and applications
  - integrate middleware with apps
  - simplify deployment
- Evolve into fully functioning scalable distributed tiered grid

http://heppc1.uta.edu/atlas/grid-testbed/index.htm



Grid Testbed Sites



US -ATLAS testbed launched February 2001

Testbed Fabric

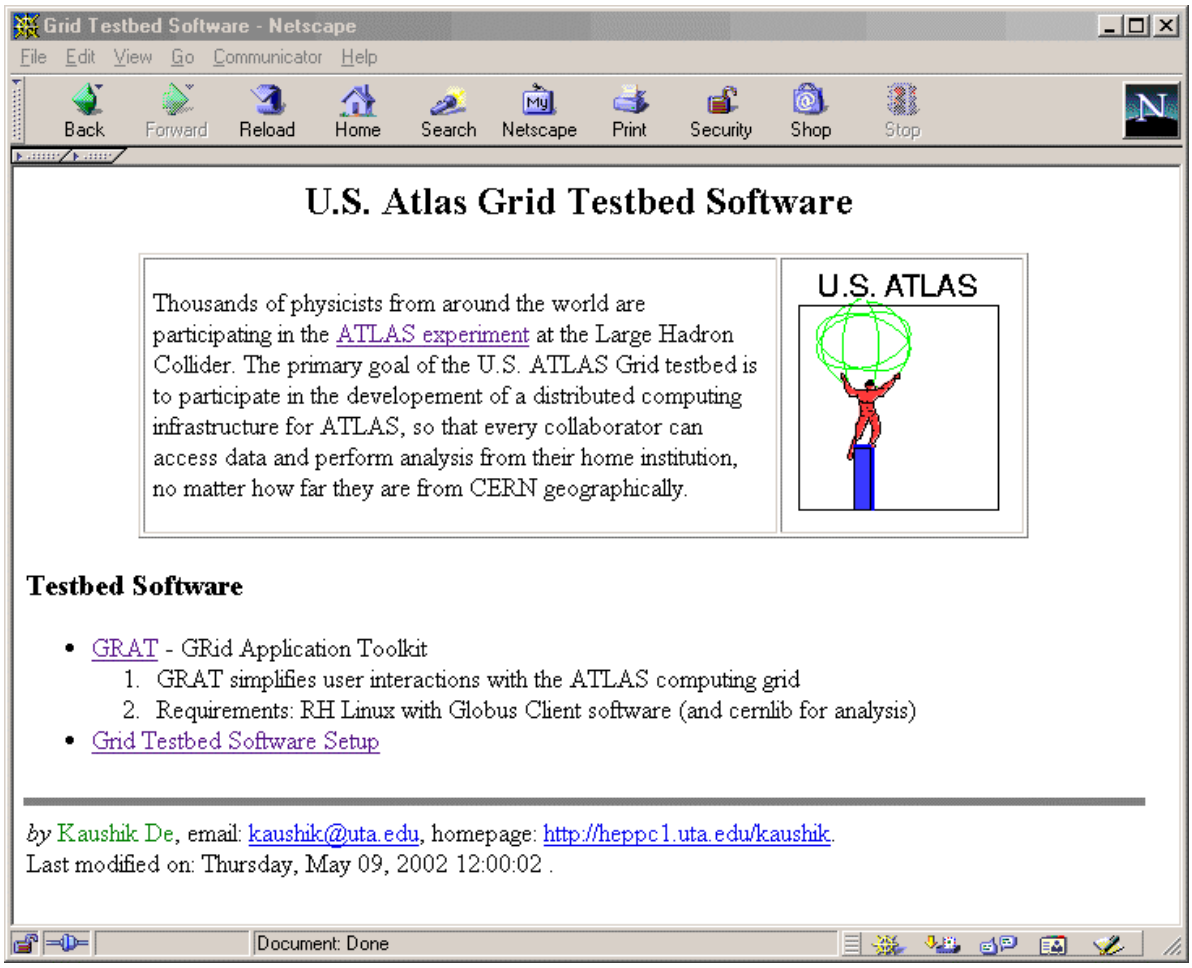
- 8 gatekeepers - ANL, BNL, LBNL, BU, IU, UM, OU, UTA
- Farms - BNL, LBNL, IU, UTA
- + Multiple R&D gatekeepers
  - gremlin@bnl - iVDGL GIIS
  - heppc5@uta - ATLAS GIIS
  - atlas10/14@anl - EDG testing
  - heppc6@uta+bnl - glue schema
  - heppc17/19@uta - GRAT devel
  - @iu - Grappa portal
  - ? - iVDGL-2 testbed

Testbed History

- Testbed continuously in operation since February, 2001
- During first year - 'proof of grid principle' tests (but mostly waiting for software ...)
- Many tools developed - GridView, Magda, Gripe, Pacman, Grappa...
- Recently: plan for deployment
  - Software distribution to sites
  - Application packaging for grid
  - Data production & cataloguing
  - User analysis tools

Software Packaging

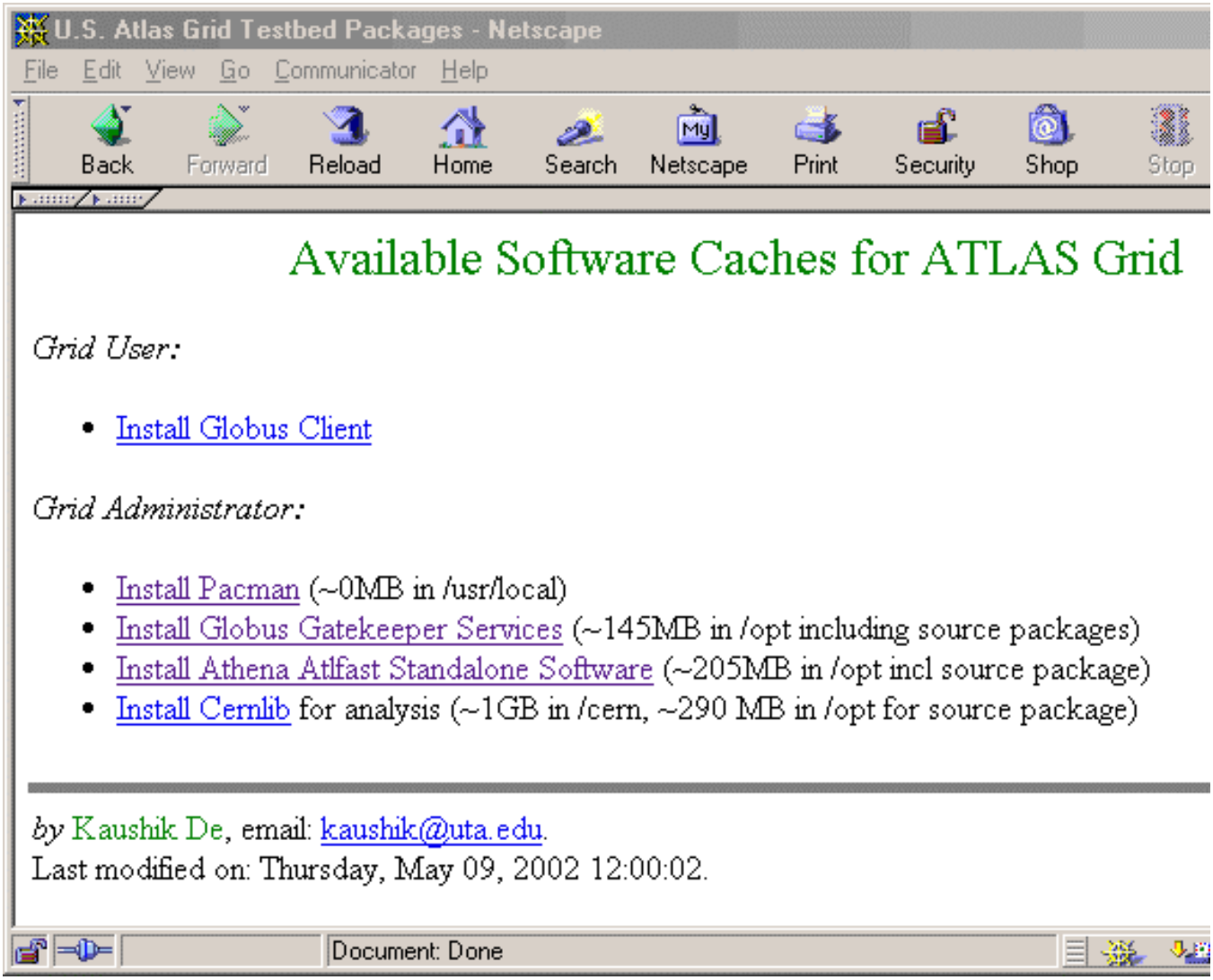
- Goals:
  - Easy installation by Sys Admins
  - Uniform software versions
- Pacman perfect for this task



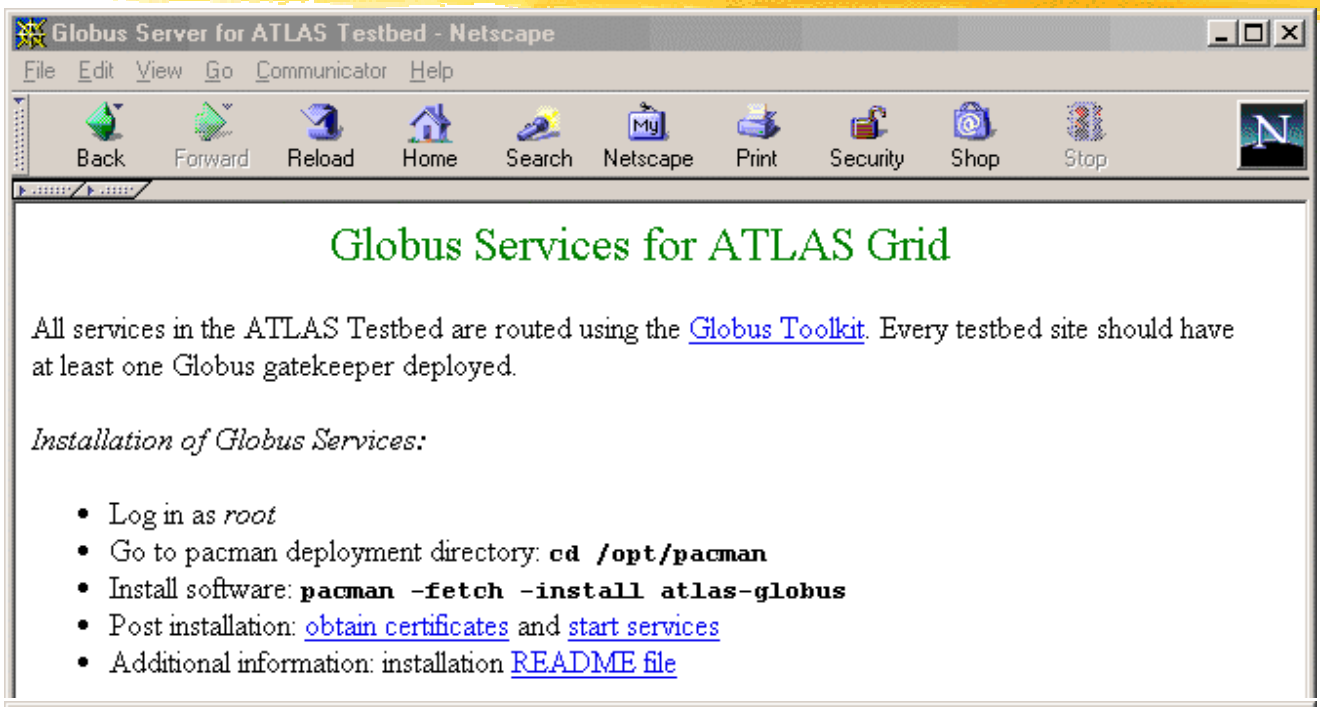
Software Distribution

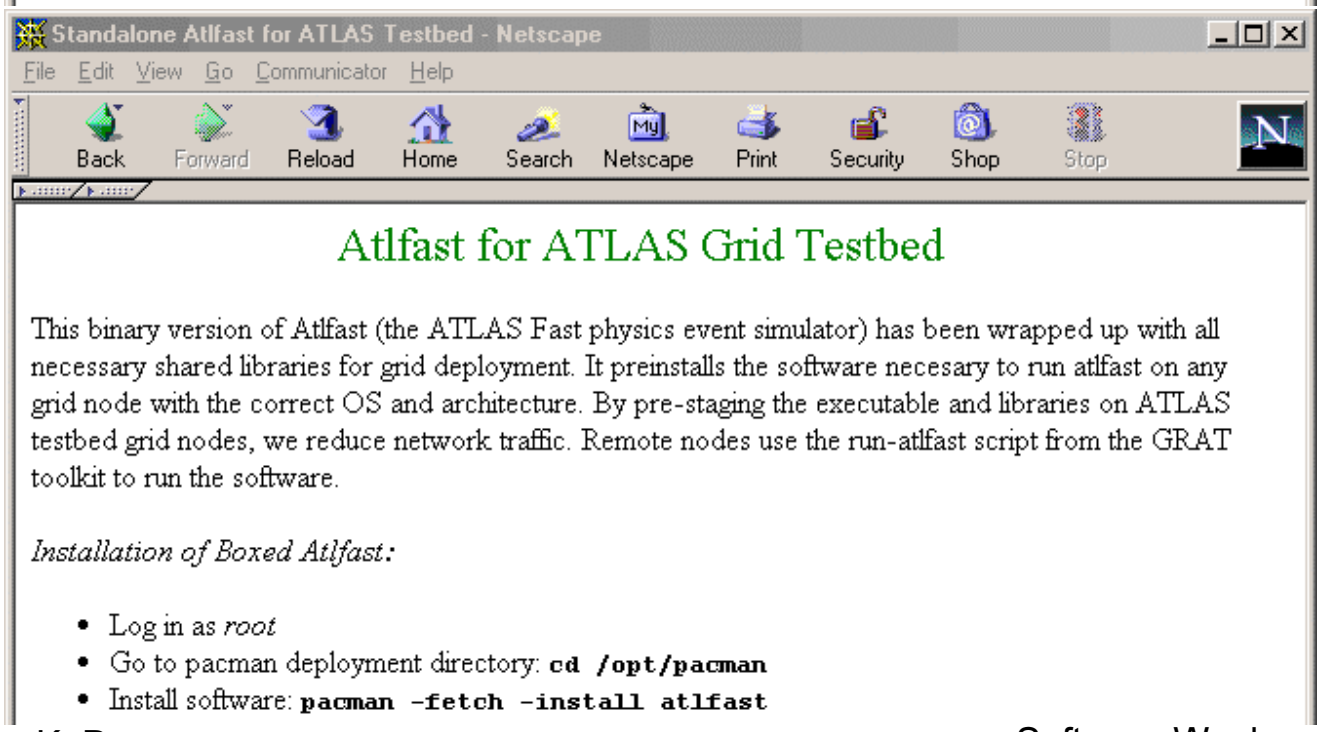
- Jason Smith, KD, Saul Youssef, Ed May, Eric Myers
- Uniform OS through kickstart
- Running RH 7.2 √
- First stage deployment
  - Pacman, Globus 2.0b, cernlib √
  - Simple application package √
- Second stage deployment
  - Magda, Chimera, VDT…
- Third stage
  - MC production, Grappa…

Available Packages



Package Examples





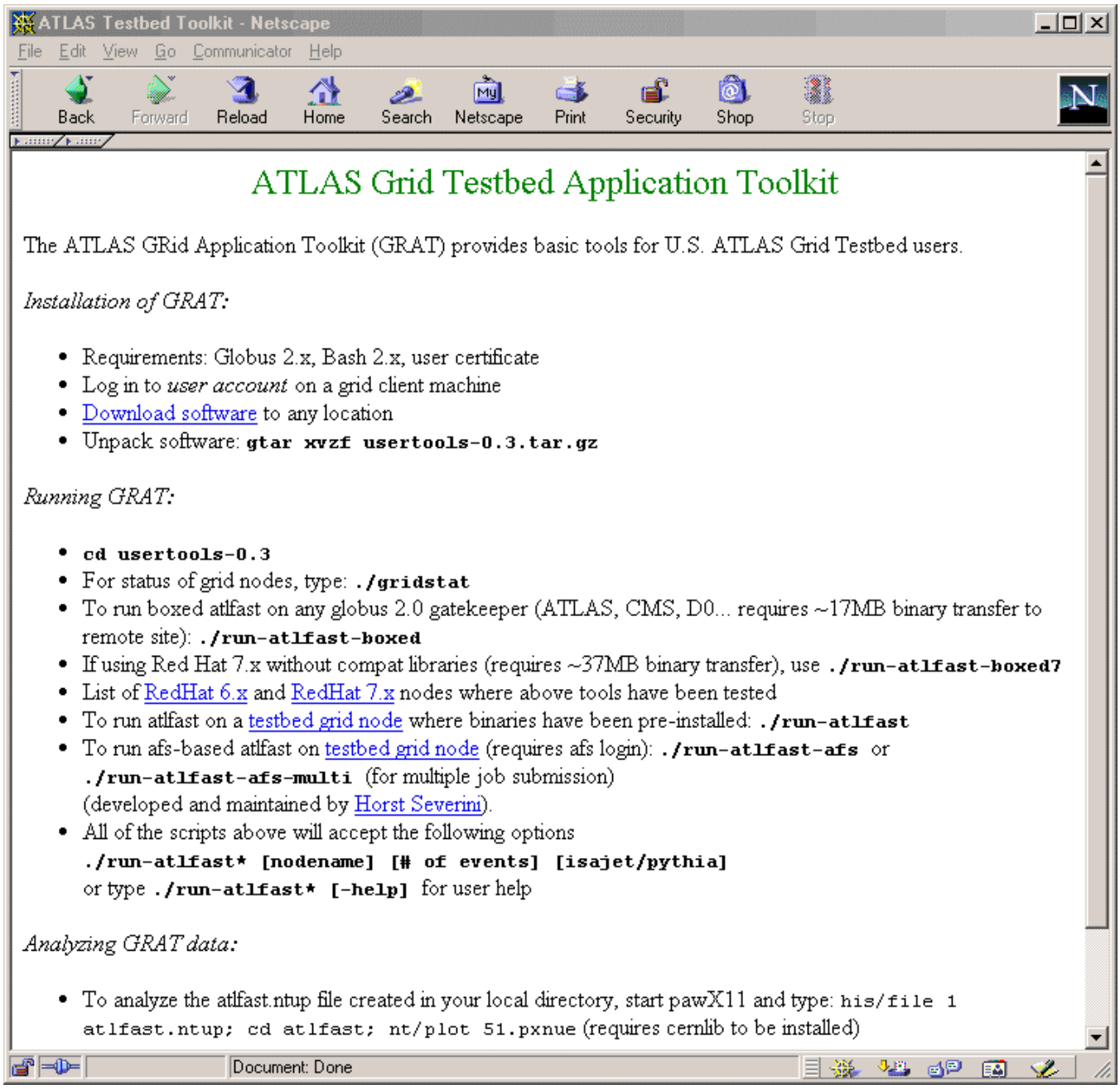
Applications Team

- Horst Severini, KD, Ed May, Wensheng Deng…
- Athena-atlfast boxed for grid (based on Julian Phillips' work)
- GRid Applications Toolkit: GRAT
  - version 0.1 released 4/12/02 √
  - tested successfully on 17 U.S. ATLAS gatekeepers, CMS gatekeeper, D0 gatekeeper, EDG CE node (RH 6.x and RH 7.x)
- Next, add basic jobmanager functions using Magda+VDC+…

GRAT Basics

- Goals:
  - do physics (without worrying about underlying middleware or ATLAS software)
- Athena-Atlfast for grid testbed
  - Tool 1: runs on any globus enabled node (requires transfer of ~17MB package)
  - Tool 2: runs on grid site where packaged software has been preinstalled
  - Tool 3: runs on afs enabled sites (latest version of software is built/used)
- Other tools
  - to check status of grid nodes
  - simple lightweight layered jobmanager (scheduler, data catalogue, file replication, virtual data …)
- Version 0.3 of GRAT: May 8, 2002

GRAT v 0.3


ATLAS Testbed Toolkit - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop Stop
ATLAS Grid Testbed Application Toolkit
The ATLAS GRid Application Toolkit (GRAT) provides basic tools for U.S. ATLAS Grid Testbed users.
Installation of GRAT:
Requirements: Globus 2.x, Bash 2.x, user certificate
Log in to user account on a grid client machine
Download software to any location
Unpack software: gtar xvzf usertools-0.3.tar.gz
Running GRAT:
cd usertools-0.3
For status of grid nodes, type: ./gridstat
To run boxed atlfast on any globus 2.0 gatekeeper (ATLAS, CMS, D0... requires ~17MB binary transfer to remote site): ./run-atlfast-boxed
If using Red Hat 7.x without compat libraries (requires ~37MB binary transfer), use ./run-atlfast-boxed7
List of RedHat 6.x and RedHat 7.x nodes where above tools have been tested
To run atlfast on a testbed grid node where binaries have been pre-installed: ./run-atlfast
To run afs-based atlfast on testbed grid node (requires afs login): ./run-atlfast-afs or ./run-atlfast-afs-multi (for multiple job submission)
(developed and maintained by Horst Severini).
All of the scripts above will accept the following options
./run-atlfast* [nodename] [# of events] [isajet/pythia]
or type ./run-atlfast* [-help] for user help
Analyzing GRAT data:
To analyze the atlfast.ntup file created in your local directory, start pawX11 and type: his/file 1 atlfast.ntup; cd atlfast; nt/plot 51.pxnue (requires cernlib to be installed)
Document: Done

MC Production

- MC production traditionally uses dedicated CPU farms
- Grid - a new paradigm for distributed MC production
- Testbed production plans:
  - Short production sequence to test middleware and tools, based on GRAT/Grappa athena-atlfast
  - Long production sequence to test fabric capabilities based on DC1 toolkit

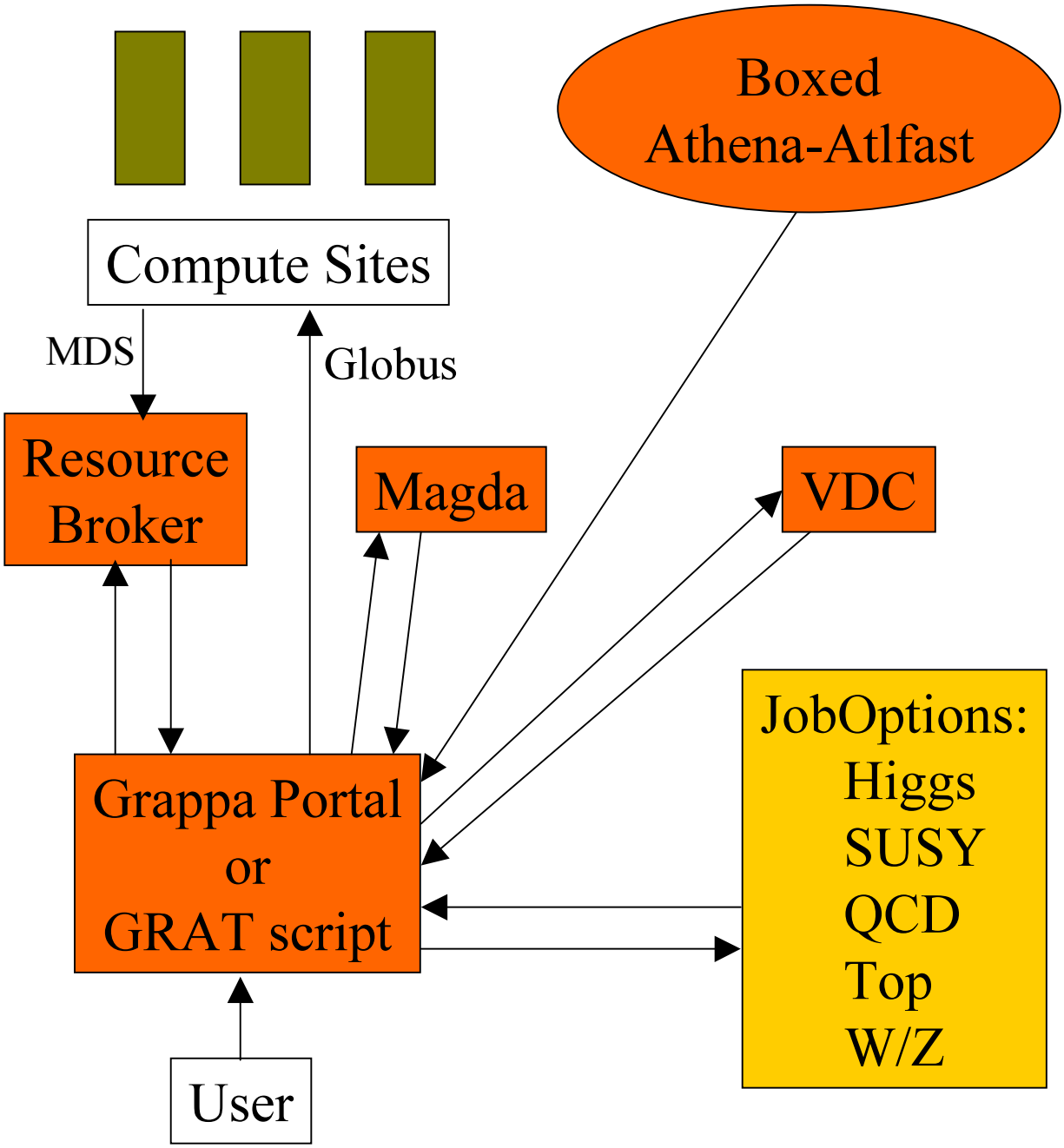
Long Production

- Currently limited U.S. participation in DC1 - BNL only
- Being re-packaged for U.S. grid deployment
- Distribute to all testbed sites
- CPU intensive - use farms
- Quality of service important
- Emphasis on Virtual Data
- Plan: June-November 2002

Short Production

- Application: Athena-atlfast-boxed
- Middleware: Globus+Magda+VDC
- Interface: GRAT, Grappa
- Production: 8 ATLAS testbed sites + 2 CMS testbed sites + 2 D0 MC farms + ? EDG sites
- Data:
  - Phase 0: 10^7 events
  - Phase 1: 10^8 events
  - Phase 2: 10^9 events, 1 TB storage, 40k cataloged files
- Schedule: June-October 2002

Short Production Architecture



Integration

- EDG - Ed May, Jerry Gieraltowski
- Virtual Data - Sasha Vaniachin, Pavel Nevski, David Adams
- Afs Software - Alex Undrus, Shane Canon, Iwona Sakrejda
- Networking - Shawn McKee, Rob Gardner
- Data management - Wensheng Deng, Ed May, David Malon
- CAS, ...

Monitoring

- Dantong Yu, Patrick McGuigan, Craig Tull, KD, Dan Engh
- Site monitoring - MDS
  - publish BNL acas information √
  - Glue schema testbed √
- Software installation - MDS
  - pacman information provider
- Application monitoring
  - instrument athena
- Grid monitoring - MDS+visual
  - GridView, Ganglia…
  - hierarchical GIIS server √

Testbed Tools

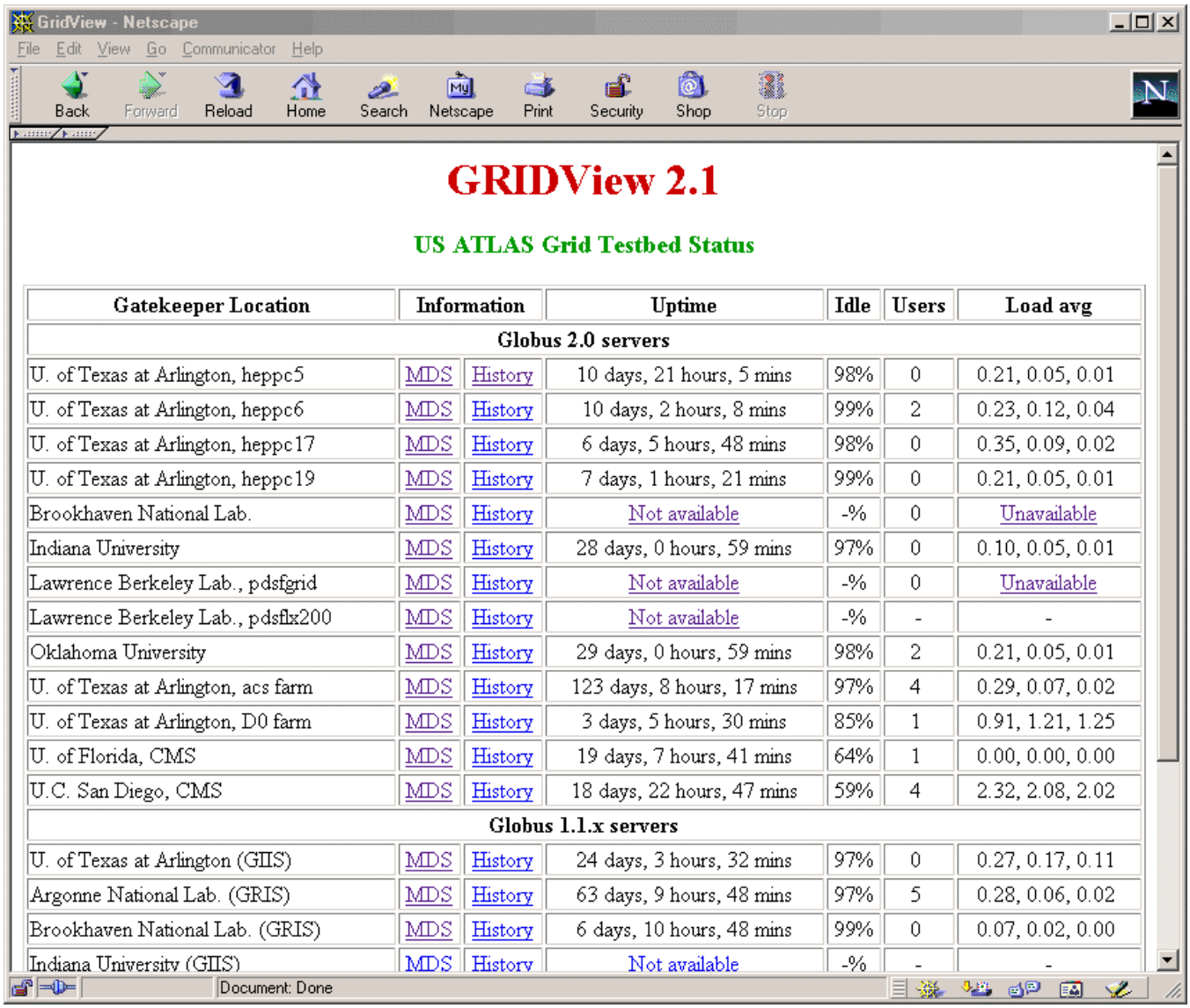
- Many tools developed in U.S. ATLAS testbed during past year
- GridView - simple tool to monitor status of testbed (Java version being developed) Patrick McGuigan
- Gripe - unified user accounts Rob Gardner
- Magda - MAnager for Grid Data Torre Wenaus, Wensheng Deng
- Pacman - package management and distribution tool Saul Youssef
- Grappa - web portal using active notebook technology Shava Smallen
- ...

GridView Tool

- Monitoring is critically important in distributed Grid computing
  - to check system health
  - for resource discovery
  - for job scheduling and resource allocation decisions...
- GridView - a simple visualization tool using Globus Toolkit
  - First native Globus application for ATLAS grid (March 2001)
  - Collects information using Globus tools. Archival information is stored in MySQL server on a different machine. Data published through web server on a third machine.

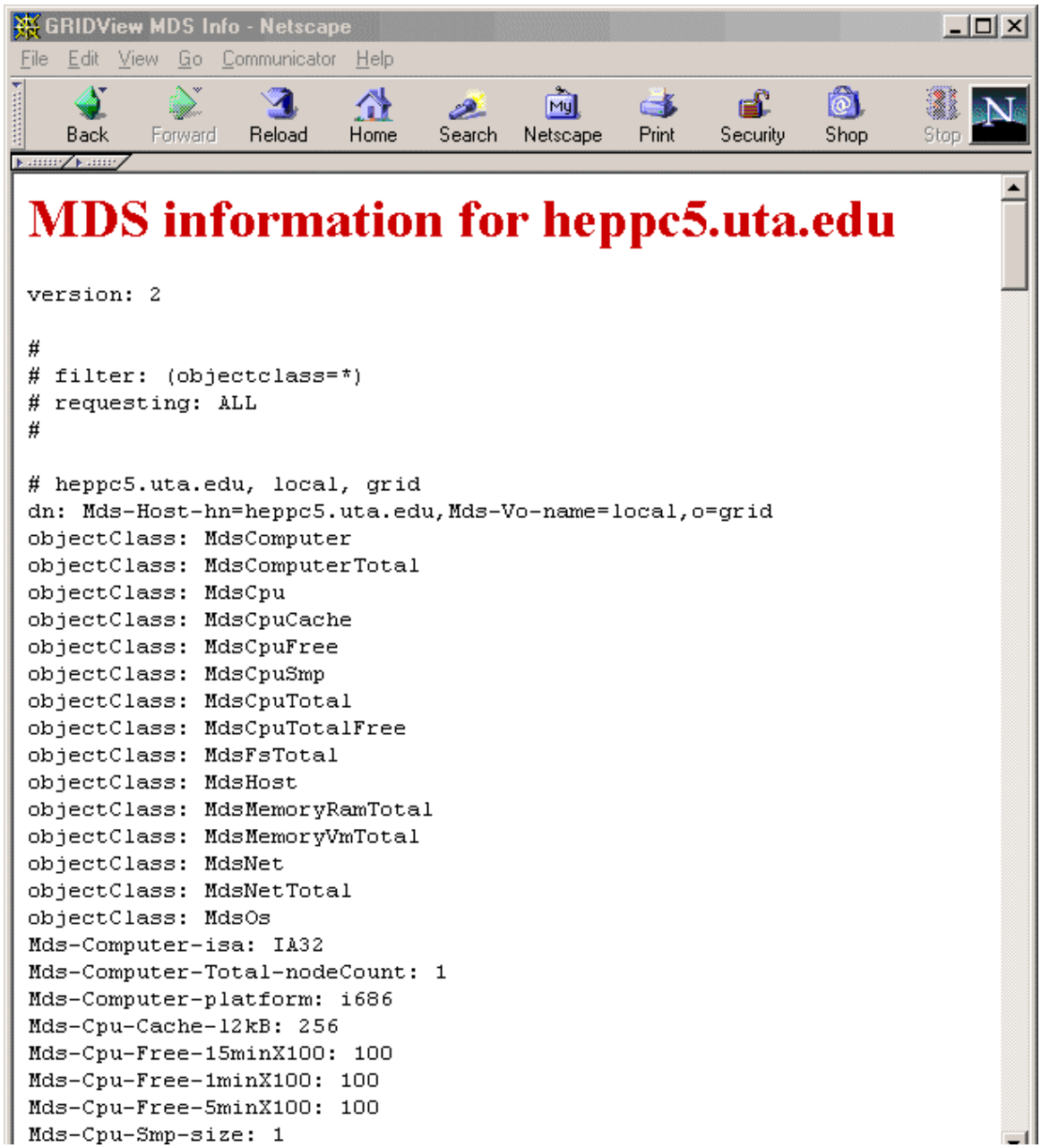
GridView 2.1

- http://heppc1.uta.edu/atlas/grid-status/index.html



Gatekeeper Location	Information		Uptime	Idle	Users	Load avg
Globus 2.0 servers						
U. of Texas at Arlington, heppc5	MDS	History	10 days, 21 hours, 5 mins	98%	0	0.21, 0.05, 0.01
U. of Texas at Arlington, heppc6	MDS	History	10 days, 2 hours, 8 mins	99%	2	0.23, 0.12, 0.04
U. of Texas at Arlington, heppc17	MDS	History	6 days, 5 hours, 48 mins	98%	0	0.35, 0.09, 0.02
U. of Texas at Arlington, heppc19	MDS	History	7 days, 1 hours, 21 mins	99%	0	0.21, 0.05, 0.01
Brookhaven National Lab.	MDS	History	Not available	-%	0	Unavailable
Indiana University	MDS	History	28 days, 0 hours, 59 mins	97%	0	0.10, 0.05, 0.01
Lawrence Berkeley Lab., pdsfgrid	MDS	History	Not available	-%	0	Unavailable
Lawrence Berkeley Lab., pdsflx200	MDS	History	Not available	-%	-	-
Oklahoma University	MDS	History	29 days, 0 hours, 59 mins	98%	2	0.21, 0.05, 0.01
U. of Texas at Arlington, acs farm	MDS	History	123 days, 8 hours, 17 mins	97%	4	0.29, 0.07, 0.02
U. of Texas at Arlington, D0 farm	MDS	History	3 days, 5 hours, 30 mins	85%	1	0.91, 1.21, 1.25
U. of Florida, CMS	MDS	History	19 days, 7 hours, 41 mins	64%	1	0.00, 0.00, 0.00
U.C. San Diego, CMS	MDS	History	18 days, 22 hours, 47 mins	59%	4	2.32, 2.08, 2.02
Globus 1.1.x servers						
U. of Texas at Arlington (GIIS)	MDS	History	24 days, 3 hours, 32 mins	97%	0	0.27, 0.17, 0.11
Argonne National Lab. (GRIS)	MDS	History	63 days, 9 hours, 48 mins	97%	5	0.28, 0.06, 0.02
Brookhaven National Lab. (GRIS)	MDS	History	6 days, 10 hours, 48 mins	99%	0	0.07, 0.02, 0.00
Indiana University (GIIS)	MDS	History	Not available	-%	-	-

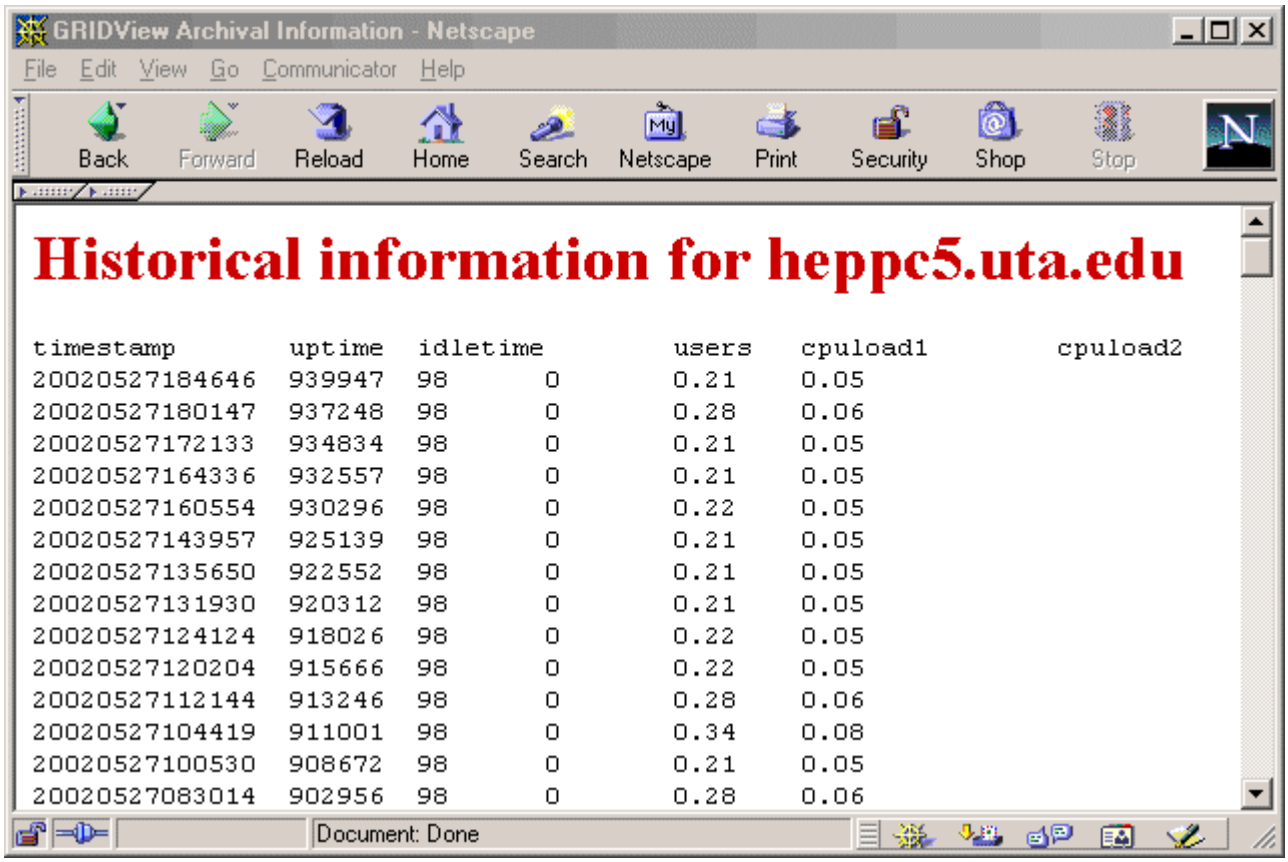
GridView MDS Info



Listing of available object classes

GridView Archive

- 6 months of history available in MySql database
- Visual tools being developed
- Add more variables soon



GridView Future

- Patrick McGuigan is working on a Java cog based version
- Better visualization
  - Historical plots
  - Hierarchical MDS information
  - Graphical view of system health
- New MDS schemas
- Optimize archived variables
- Publishing historical information through GIIS servers??
- Explore discovery tools
- Explore scalability to large systems
- Ganglia, Cricket?

Magda Tool

- MAnager for Grid-based DAta
- Designed for 'managed production' *and* 'chaotic end-user' usage
- Designed for rapid development of components to support users quickly, with components later replaced by Grid Toolkit elements
  - Deploy as an evolving production tool and as a testing ground for Grid Toolkit components
- Adopted by ATLAS for 2002 ATLAS Data Challenges
- Developers - T. Wenaus and W. Deng (pdoc) and new hire

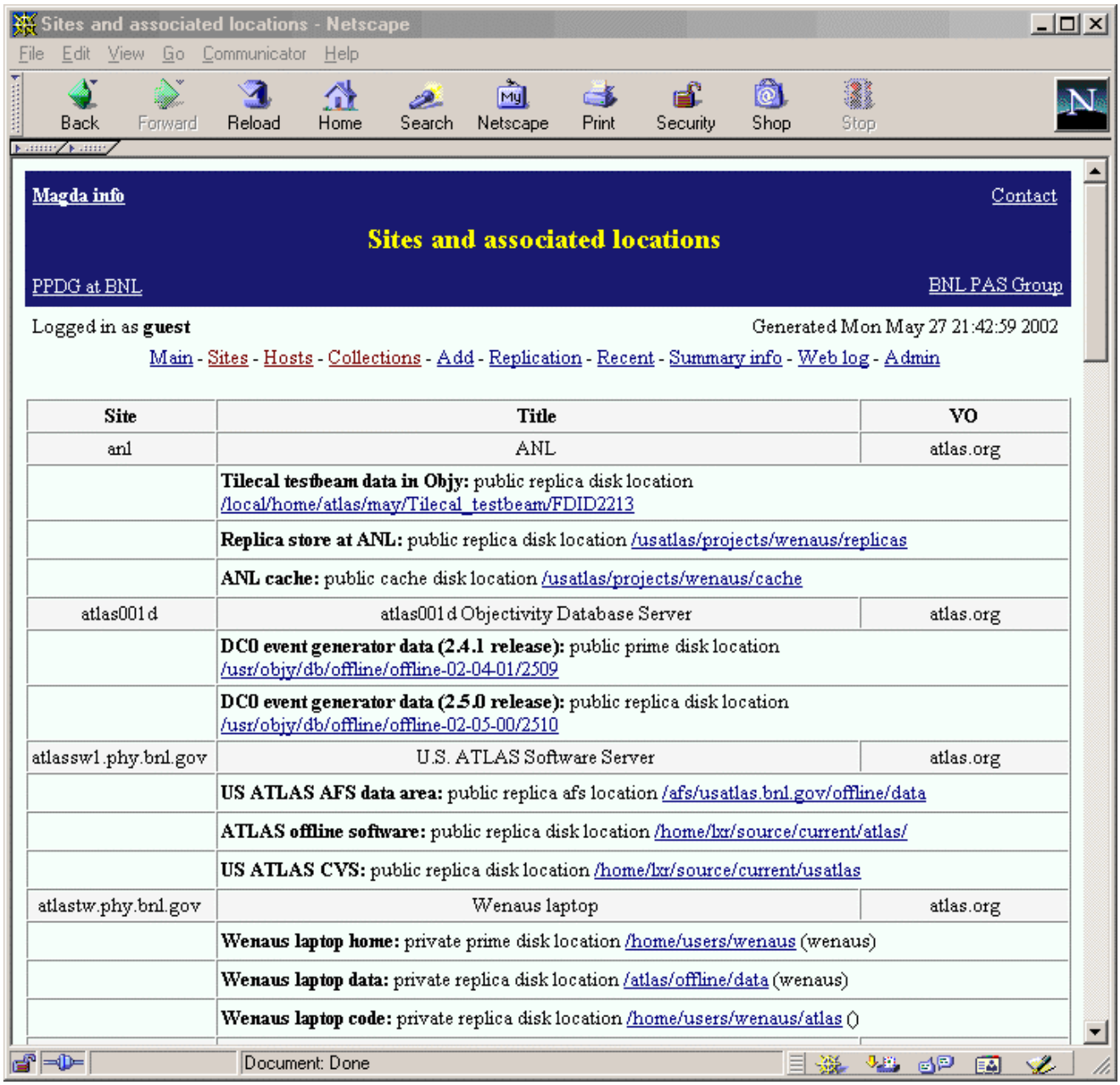
Info: http://www.usatlas.bnl.gov/magda/info

Engine: http://www.usatlas.bnl.gov/magda/dyShowMain.pl

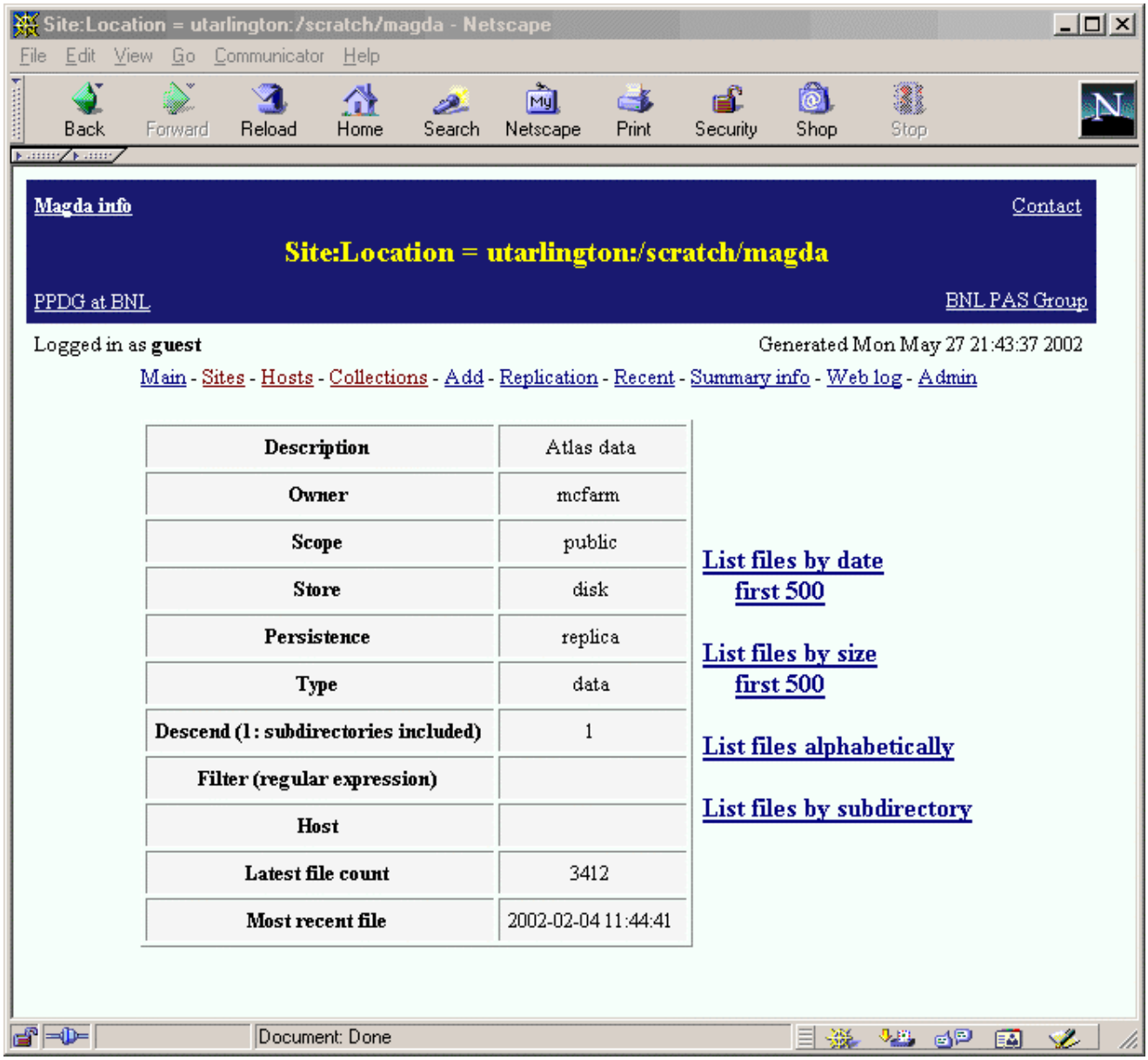
Magda Architecture & Schema

- MySQL database at the core of the system
  - DB access via perl, C++, java, cgi (perl) scripts; C++ and Java APIs auto-generated off the MySQL DB schema
- User interaction via web interface and command line
- Principal components:
  - File catalog covering any file types
  - Data repositories organized into sites, each with its locations
  - Computers with repository access: a host can access a set of sites
  - Logical files can optionally be organized into collections
  - Replication operations organized into tasks

Magda Sites


Sites and associated locations - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop Stop
Magda info
Contact
Sites and associated locations
PPDG at BNL
BNL PAS Group
Logged in as guest
Generated Mon May 27 21:42:59 2002
Main - Sites - Hosts - Collections - Add - Replication - Recent - Summary info - Web log - Admin
Site | Title | VO
anl | ANL | atlas.org
Tilecal testbeam data in Objy: public replica disk location /local/home/atlas/may/Tilecal_testbeam/FDID2213
Replica store at ANL: public replica disk location /usatlas/projects/wenaus/replicas
ANL cache: public cache disk location /usatlas/projects/wenaus/cache
atlas001d | atlas001d Objectivity Database Server | atlas.org
DC0 event generator data (2.4.1 release): public prime disk location /usr/objy/db/offline/offline-02-04-01/2509
DC0 event generator data (2.5.0 release): public replica disk location /usr/objy/db/offline/offline-02-05-00/2510
atlassw1.phy.bnl.gov | U.S. ATLAS Software Server | atlas.org
US ATLAS AFS data area: public replica afs location /afs/usatlas.bnl.gov/offline/data
ATLAS offline software: public replica disk location /home/lxr/source/current/atlas/
US ATLAS CVS: public replica disk location /home/lxr/source/current/usatlas
atlastw.phy.bnl.gov | Wenaus laptop | atlas.org
Wenaus laptop home: private prime disk location /home/users/wenaus (wenaus)
Wenaus laptop data: private replica disk location /atlas/offline/data (wenaus)
Wenaus laptop code: private replica disk location /home/users/wenaus/atlas ()
Document: Done

Magda Actions

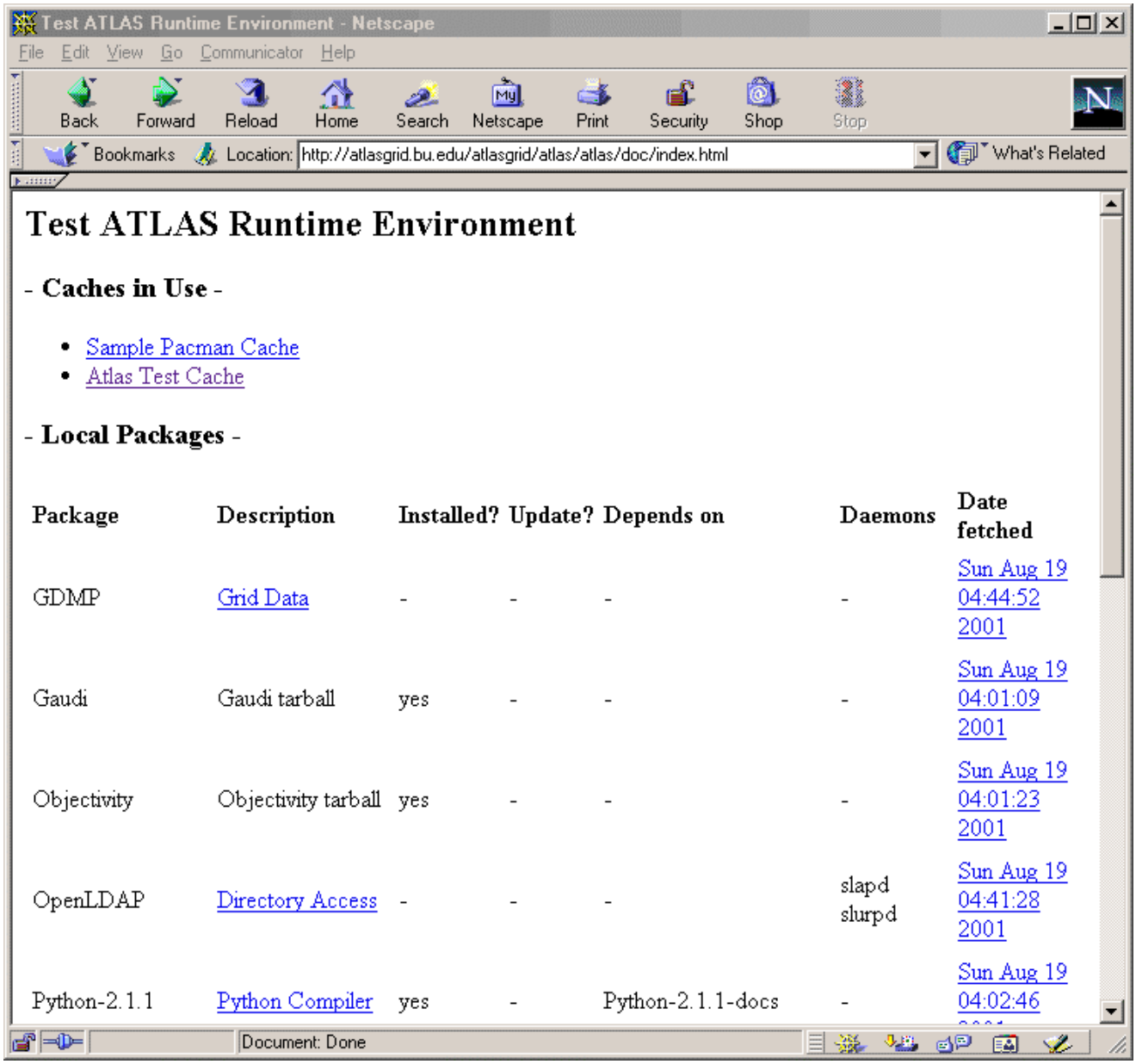


Pacman

- Procedures for fetching, installing and setting up software can be figured out once and exported via caches which everyone then uses.
- For cache users, fetching, installation, setup and updates are automatic.
- Making caches is easy and doesn't involve repackaging or even keeping separate copies of the software.
- Packages which depend on other packages are handled correctly.
- If something breaks you complain to the cache manager instead of fixing it yourself.

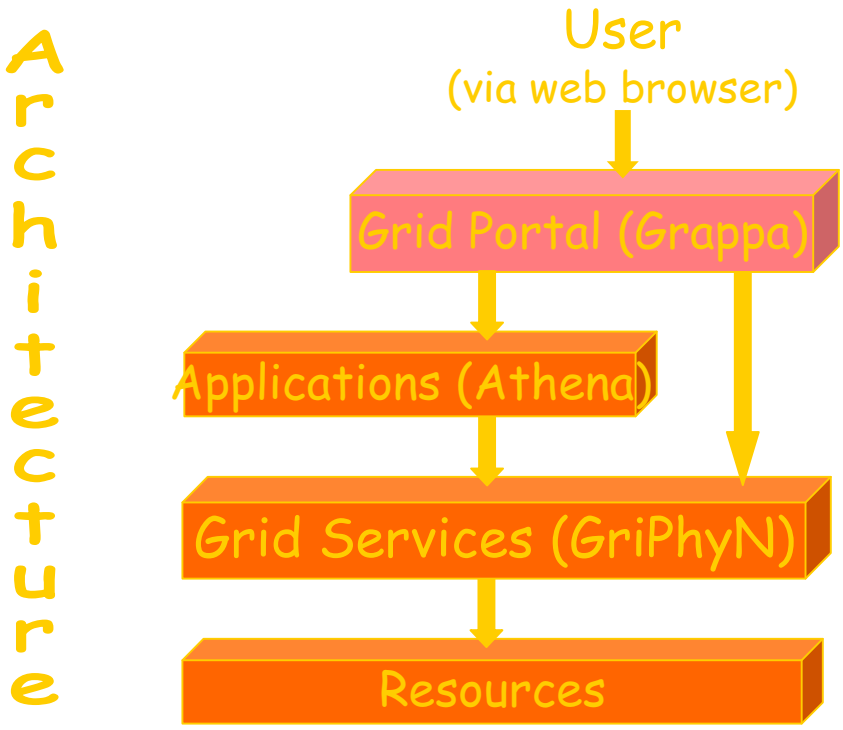
```
% pacman -fetch -install
atlas_grid_testbed
```

Pacman Example



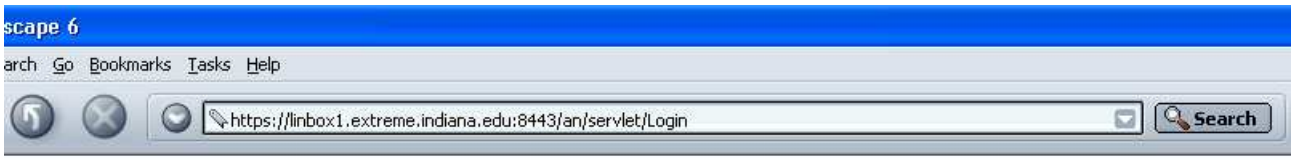
Grappa

- Lots of GriPhyN/Grid services
- Higher-level user interface
- Use a Grid portal



- Goal: provide a Grid portal interface by which physicists can launch and manage jobs and data

Grappa Login



Welcome to the XCAT Science Portal!

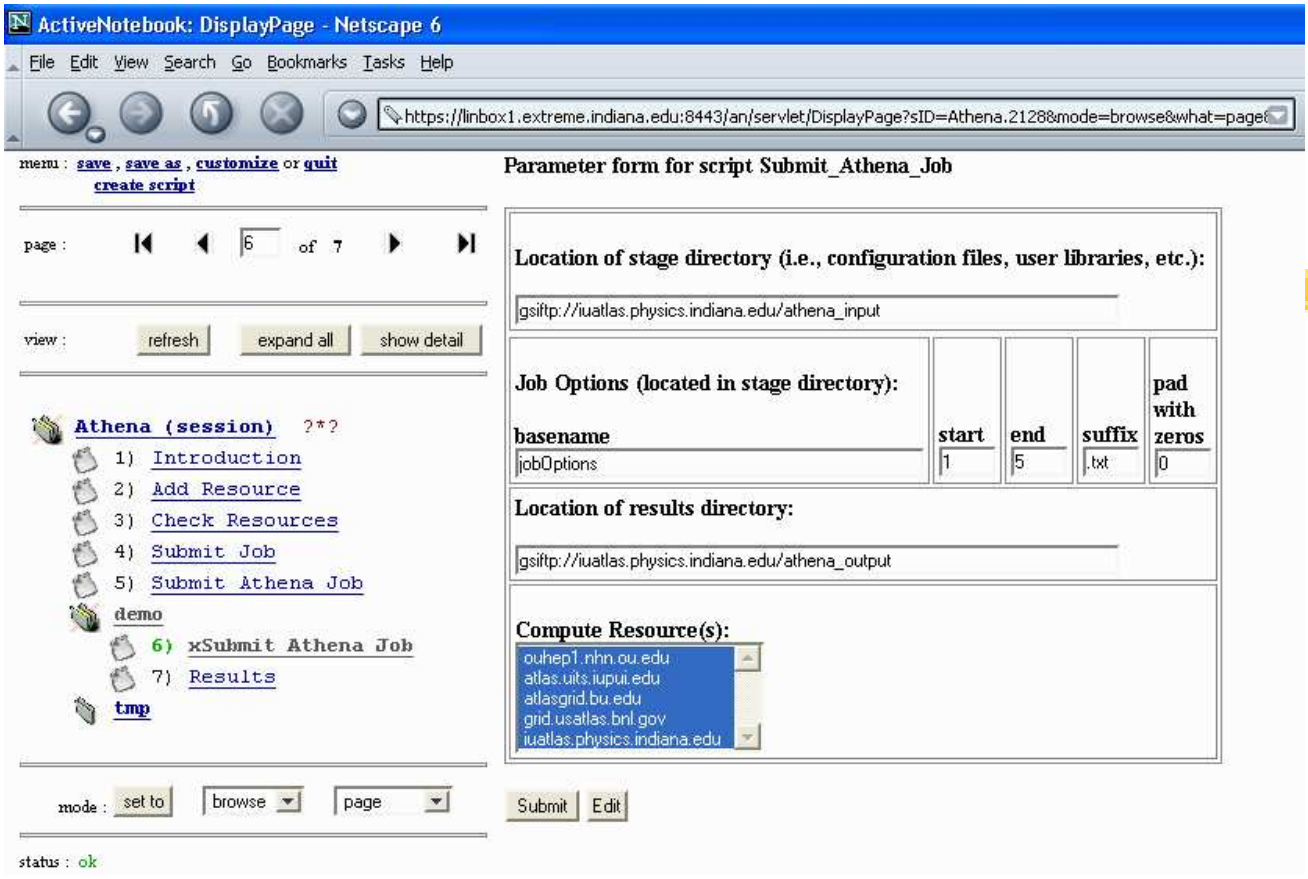
Please enter your portal username and GSI passphrase
to log into the XCAT Science Portal system © :

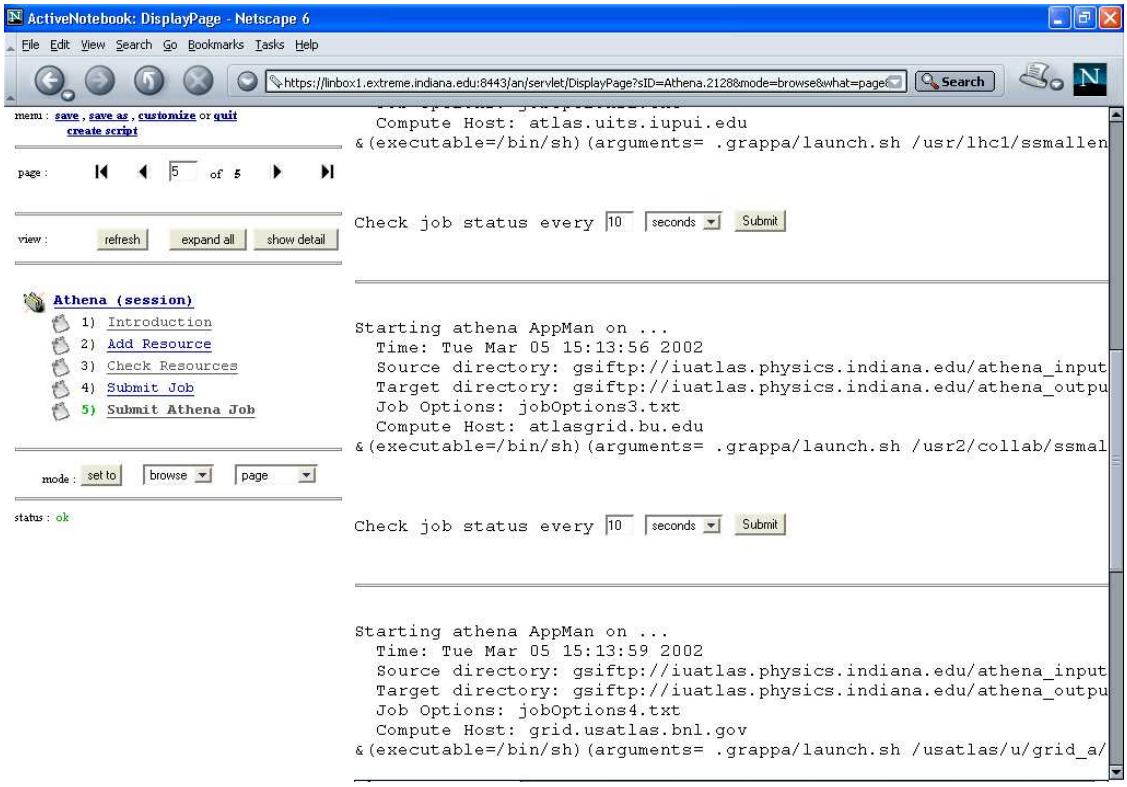
Portal Username:* ssmallen

GSI Passphrase: *******

login

* Note: If you don't have a portal username, it will be created the first time you log in.


ActiveNotebook: DisplayPage - Netscape 6
File Edit View Search Go Bookmarks Tasks Help
https://linbox1.extreme.indiana.edu:8443/an/servlet/DisplayPage?sID=Athena.2128&mode=browse&what=page
menu : save , save as , customize or quit
create script
page : 6 of 7
view : refresh expand all show detail
Athena (session) ?*?
1) Introduction
2) Add Resource
3) Check Resources
4) Submit Job
5) Submit Athena Job
demo
6) xSubmit Athena Job
7) Results
tmp
mode : set to browse page
status : ok
Parameter form for script Submit_Athena_Job
Location of stage directory (i.e., configuration files, user libraries, etc.):
gsiftp://iuatlas.physics.indiana.edu/athena_input
Job Options (located in stage directory):
basename
jobOptions
start
1
end
5
suffix
.txt
pad with zeros
0
Location of results directory:
gsiftp://iuatlas.physics.indiana.edu/athena_output
Compute Resource(s):
ouhep1.nhn.ou.edu
atlas.uits.iupui.edu
atlasgrid.bu.edu
grid.usatlas.bnl.gov
iuatlas.physics.indiana.edu
Submit Edit


ActiveNotebook: DisplayPage - Netscape 6
Compute Host: atlas.uits.iupui.edu
&(executable=/bin/sh)(arguments= .grappa/launch.sh /usr/lhc1/ssmallen
Check job status every 10 seconds Submit
Starting athena AppMan on ...
Time: Tue Mar 05 15:13:56 2002
Source directory: gsiftp://iuatlas.physics.indiana.edu/athena_input
Target directory: gsiftp://iuatlas.physics.indiana.edu/athena_outpu
Job Options: jobOptions3.txt
Compute Host: atlasgrid.bu.edu
&(executable=/bin/sh)(arguments= .grappa/launch.sh /usr2/collab/ssmal
Check job status every 10 seconds Submit
Starting athena AppMan on ...
Time: Tue Mar 05 15:13:59 2002
Source directory: gsiftp://iuatlas.physics.indiana.edu/athena_input
Target directory: gsiftp://iuatlas.physics.indiana.edu/athena_outpu
Job Options: jobOptions4.txt
Compute Host: grid.usatlas.bnl.gov
&(executable=/bin/sh)(arguments= .grappa/launch.sh /usatlas/u/grid_a/

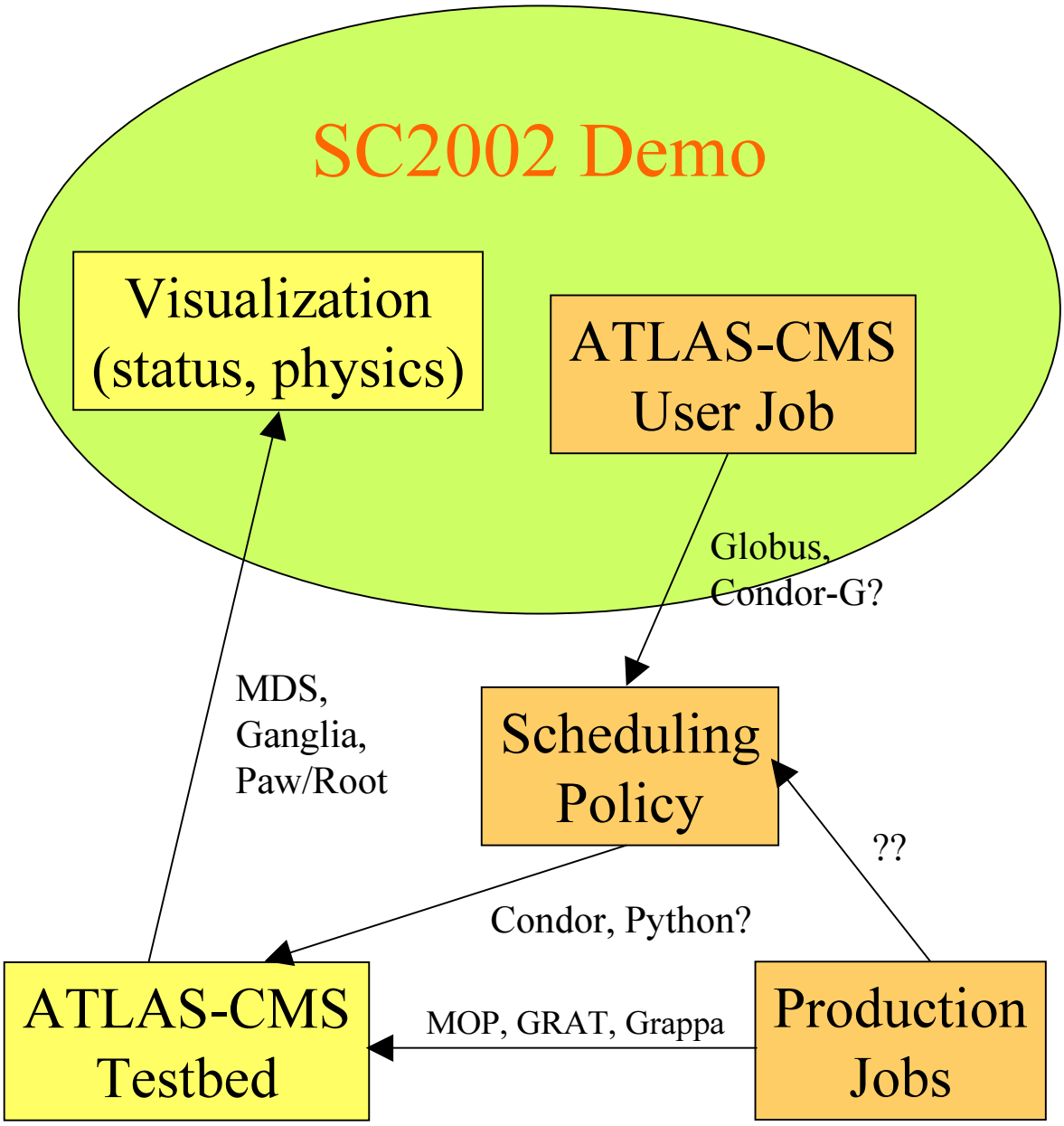
SC2002 Plans

- Supercomputing 2002 in November
- ATLAS-CMS demo (FNAL/SLAC)
  - Rick Cavanaugh & KD
  - preliminary schedule agreed
  - GRAT already running at CMS sites
- Production demo 1 (BNL)
  - conceptual plan agreed
  - Monitor long & short grid production
- Production demo 2
  - plan under discussion
  - Virtual data and Magda
- Application monitoring (LBNL)
  - Athena + Netlogger + Prophesy
- Grappa demo (IU/ANL)

SC2002 ATLAS-CMS Demo. Schedule

- Agreed to by ATLAS-CMS 4/26/02
  - 1 week - identify team from CMS & ATLAS
  - 3-4 weeks - demonstrate CMS simulation working on ATLAS testbed and vice versa
  - 6-8 weeks - develop policy scheme and implementation plan
  - July - meet to evaluate progress and finalize demo
  - August - first version of visual demo tools
  - September - first real ATLAS-CMS joint run
  - October - polish final demo

Demo. parts


SC2002 Demo
Visualization (status, physics)
ATLAS-CMS User Job
Globus, Condor-G?
MDS, Ganglia, Paw/Root
Scheduling Policy
??
Condor, Python?
ATLAS-CMS Testbed
MOP, GRAT, Grappa
Production Jobs